Geoffrey Fox (Indiana University. See talk *2020-2025 Scientific Computing Environments (Distributed Computing in an Exascale era)*)

Fox introduced the presentations with two broad questions and some subtopics to consider.

Questions
1. The components of national research computing in exascale era with mix of high end machines, clouds (whatever commercial companies offer broadly or publicly), university centers, high throughput systems and with growing amounts of distributed and "repositorized" data serving High End and Long Tail researchers.
2. The nature of an environment like XSEDE in the Exascale era; i.e. the nature of a distributed system of facilities including one or more exascale machines. Should it be relatively tightly coupled like XSEDE or more loosely coupled like DoE leadership systems (or both!)

SubTopics
1. A) What services do 2025 science projects need from cyberinfrastructure; examples are -- Collaboration; On-demand computing; Digital observatory; High-speed scratch and persistent storage; Data preservation; Identity, profile, group management, reproducibility of results, versioning, and documentation of results
2. B) What are requirements in 2025 -- are there changes in distributed system requirements outside details of exascale machines with their novel architecture e.g.
   - Will big data lead to new requirements
   - Will feeding/supporting exascale machine lead to new requirements
   - Will supporting long tail of science lead to new requirements
   - Can we do more to make people use central services rather than building their own

Miron Livny (Wisconsin)
The important issues behind distributed computing have not changed since the work of P H Enslow of Georgia Tech in 1970's but even now, we are still correcting flaws in previous work. Most of speakers today are "sellers of technology" with domain scientists as "customers". We need to spend more time talking to users and correspondingly less on new technology. In future we expect more financial stress but need to support people that are greater in number and variety.

There are still fundamental gaps. We are OK on computing but still poor on storage. Recently Livny was in a meeting discussing such basic issues as how to decide what data to keep and what to throw away within a fixed budget. There will always be debates as to whether to use clouds, laptops, smartphones etc. for ones computing. However we will always need resource management. Software Defined Networks will be powerful but add yet another thing for us to manage.

Livny does not have a clear plan for the exascale time frame but is sure that we should work more on problems and less on technology silver bullets. We will need to understand "Long Tail" scientists better. How much complexity should be exposed. The exponential changes requires us to rethink approaches. We will need to help scientists produce the computational tools they need to produce their solutions even if this looks like engineering at times. Perhaps a new venue is needed to discuss the open issues here.

The Galaxy http://galaxyproject.org/ web-based platform for data intensive biomedical research is an example of a domain specific success.

Shantenu Jha (Rutgers. See talk *"Extreme-Scale" Distributed Computing Scientific Computing Environments 2025*)
We have good and continuing progress but still it is hard for individual "non-heroic" researchers to do relatively simple things such as running O(100) tasks each of O(10GB) over O(10) nodes. We are missing abstractions that enable precise reasoning and manage complexity for such questions as Linpack performance on a distributed system. Jha's P* system allows reasoning about Pilot jobs.

Comparing 2025 to now, Jha expects similar functionality to today but the heterogeneity and scale of systems and problems will increase. Jha then reviewed three examples: ATLAS (major LHC experiment), Square Kilometre Array (coming online in the 2020-25 timescale) and the European Human Brain project.

We need to separate capability from the technology used to provide functionality; this is where abstractions come in. Jha stressed need for federation of dynamic heterogeneous systems such as that between OSG and XSEDE.

A question asked if the 2025 environment will be revolutionary or evolutionary and some of both was the answer

Dennis Gannon (Microsoft Research. See talk *2025 Scientific computing environments*)
The data explosion is transforming science and leading to increasing importance of the "fourth paradigm". Fields like particle physics and astronomy understand what to do and genomics is successfully becoming a "big data big science". However there is the long tail with fields like economics and social science which need help. They don't want to be system administrators; they want to modify a few parameters in their favorite scripting language.

In 2014-2015 we will see an amazing deployment of commercial cloud infrastructure with 10 million new servers and 10 exabytes of new storage.

An exciting area of growth is machine learning and neural nets for applications like genome-wide association studies, computer vision, automatic speech translation, medical and environmental image analysis. (see slide 10 of Fox's talk). Urban science is thriving from these developments.

In 2025 Long Tail science will benefit from data being available to everybody from the cloud. Wolfram Alpha and Apple Siri are interesting models. We will see growing use of Social Network tools for scientific collaboration. We need to simplify the front end and inevitably use complex back ends; note Siri is simultaneously serving millions of users and never crashes. We need a new generation of application design methodology.

Questions discussed the use of either academic or commercial resources to support (long tail) science. Interfaces at both the Gateway (portal) level and below this are needed.

Ioan Raicu (IIT. See talk *Supporting Data-Intensive Distributed Computing in an Exascale Era*) Raicu started in this area in 2003, a similar time to Shantenu(2001) and noted his discussion was more technology centric. The first slides stressed that we are seeing an increasing number of "general purpose" cores per CPU with up to 1000 being expected by 2020. (GPU's already have more special purpose cores). This timescale will be notable for increased concurrency everywhere.

Raicu noted problems with the fragility of MPI in an era when fault tolerance will be harder. Extrapolating from today one finds that at around million nodes, the mean time between failure becomes comparable to checkpointing time. Solutions involve changing storage architecture which Raicu discussed in detail. He expects that storage and computing will come more integrated (as in Hadoop or Google file systems) and perhaps every node will have an SSD. In such a world, storage on network will be used for long term archiving only. He emphasized that tomorrow's many core chips will be able to dedicate cores to storage and network handling so such nodes will support traditional CPU dominant computations without execution time fluctuation (that lower MPI efficiency) from handling SDN's, communication and storage. We need to learn how to use HPC network topologies more efficiently to handle diverse loads. Machines like the current Argonne Blue Gene P have excellent compute networks but few storage nodes and inadequate "storage" bandwidth to them.

Decentralization with support of needed locality is critical. For example, Storage and Computing will be distributed but they will be co-located. Slide 11 of Raicu's talk lists several important topics; all of these need "distributed" added in front of them. Raicu stressed importance of NOSQL data models with distributed hash tables. These and the equally important distributed message queues are featured in many cloud PaaS platforms.

The discussion involved the application mix; will it change and will it need to have greater interactive support. Are the storage changes evolutionary or revolutionary? Raicu suggested need for efficient runtime but interfacing with loosely coupled environments; MapReduce teaches value of simplicity. Storage needs to change from traditional single CPU+Disk view.

Jim Pepin (Clemson)

He sees computation growing out of the desktop following the Long Tail theme. Further he sees the "exploding middle" -- a growing demand for middle scale computations with up to 10,000 cores.

The classic campus CIO is a network plumber focused on enterprise administrative and student support. So most campuses are not well set up to support research, In particular we need stronger links from a given campus to outside resources -- national level and other campuses. The NSF CC-NIE http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504748 program has helped and campuses are getting together in projects like the Condo of Condos http://condo-of-condos.org/. For example we need identity management that extends outside campuses. Current security models develop geographic "moats" around universities. However a chemist at one university is more likely to interact with a chemist at a different university than with a philosopher inside the same moat. CC-NIE also supports some of the engineering but not enough. We need new data management schemes and new O/S's that support the cross campus model. The NSF OCI Campus Bridging work wrote good reports https://www.nsf.gov/cise/aci/taskforces/TaskForceReport_CampusBridging.pdf; we need to follow through on implications.

ESNET introduced the interesting science DMZ concept http://fasterdata.es.net/science-dmz/ which is a step forward but a "cop-out"; the whole campus should be the DMZ.

Current networks are just too slow. Over last 35 years, maybe compute performance has increased by a factor of 10^8 but networks only by a factor 10^6. The graph in slide 5 of Fox's talk shows that network traffic is growing more slowly than Moore's law and more slowly than total storage and peak compute performance. Commercial clouds will be important for campuses but the network will still limit this.